control by more than 10 percent, the alternate hypothesis for this non-inferiority trial is that the difference of the success rate between BAK and the Charite is less than 10 percent, and non-hypothesis is that the difference of the success rate between BAK and Charite is more than delta. So the detraction of the non-hypothesis will conclude that Charite is at least as good as BAK.

Alternately, a more informative way is to construct the one-sided 95 percent confidence interval for the difference of a success rate, P sub BAK minus P sub Charite. If the upper bound of this one-sided 95 percent confidence interval is less than 10 percent delta, then we can claim non-inferiority of Charite compared to the BAK. As an example shown here, Case A, the upper bound is below the red line marked by delta. And in Case B we cannot conclude that Charite is non-inferior to BAK.

The Statistical Analysis Plan in the original IDE protocol was far from complete. Most of the patients 24 months data became available when the Statistical Analysis Plan was finalized in November

2003. The sponsors state that there is -- no income analysis was conducted and also, there is no preliminary analysis was conducted to modify the Statistical Analysis Plan.

For the primary endpoint, the overall success rate at 24 months, the primary analysis is basically a simple two group comparison of the success rate and non-inferiority hypothesis, as I mentioned in the previous slide. And also, the one-sided 95 percent confidence interval for the success rate difference between the two groups was also constructed.

And the second analysis is to evaluate the potential confounding facts from several important covariates, such as age, gender, pain medication, operative level and investigation of site, and also correlated could be added later on as needed, such as body mass index, pre-operative activity level. And also, I would like to point out there was no plan in the protocol trying to demonstrate any superiority for all the secondary components, secondary endpoints.

After randomization, a total of 205

patients were implanted with the Charite and 99 patients receiving the BAK. Overall, compared to only 79 patients in the BAK group has completed the study at 24 months without any missing data. 87 percent of patients in the Charite group had completed data at 24 months.

The non-completers with missing data at 24 months were classified into three categories, the discontinued, overdue and the not yet due patients. There were 7 percent patients in the BAK group and five patients in the Charite, 2 percent, in the Charite groups has early discontinuation. So you have noticed there is about three or more than three times more discontinued patients in the BAK compared to the Charite.

The overdue patients was defined as those patients who have not received all the components of the primary endpoint at 24 months and have not been classified as early discontinuation. And for such a population with missing data at 24 months, there is an 8 percent of BAK patients versus 5 percent Charite patients as overdue patient, and there is about the

equivalent percentage of patients, which was not yet due, because this PMA was submitted before all the randomized patients completed the 24 follow-up evaluation.

Although in the protocol the sponsor defined the ITT analysis population will be all the randomized patients, but the actual sponsor's ITT analysis include only completers and discontinued, and they treat the discontinued patients as failures, because there is a high percentage of discontinuation in the BAK compared to the Charite, so such analysis is strongly biased against the BAK in favor of the Charite. FDA believed that the true ITT analysis should include all the randomized patients with those missing data handled appropriately.

To assess the impact of missing data on the comparative evaluation of the success rate between the two groups, sensitivity analysis was conducted under several different scenarios as shown in this slide, and this slide, the bar is a 95 percent confidence inflow, a one-sided, for the difference of a success rate at 24 months.

Again, as shown in the left panel, there was a high percentage of non-completers, 21 percent in the BAK group compared to Charite, 13 percent. Therefore, any analysis excluding those non-completers or including them all failures will lead to a biased estimate in favor of Charite device.

For example, if you include only completers, all the actual sponsor's ITT analysis, which is completers plus discontinued, or all the randomized patients including non-completed as failures will be biased against the control, BAK, in favor of the Charite.

But if we include all randomized patients with missing data at 24 months and treat them as all success in favor of the BAK, the upper bound of the 95 percent confidence interval for the difference of a success rate is almost 7 percent, meaning that the Charite could be worse than BAK by almost 7 percent in terms of the success rate at 24 months. But using the non-inferiority margin, delta 10 percent, we still can claim the non-inferiority of Charite compared to BAK.

The sponsors also the last observation

carried forward to impute the missing data at 24 months for the non-completers, and FDA also looked at the details of the sponsors as LOCF and proposed a modified conservative LOCF, and that I'm going to talk about in the next three slides.

Before going there, I also would like to point out that in the worst case scenario where we treat all the non-completers as success for the BAK, but a failure for the Charite, such a conservative way in favor of the BAK, then the one-sided 95 percent confidence interval of the difference, the upper bound of that is 21 percent. It's well beyond the non-inferiority 10 percent margin. So under the worst case scenario, the Charite device will not be claimed as non-inferior to BAK.

So now let's move onto the last observation carried forward analysis. The last observation carried forward analysis carries forward the last available observation available at the last time to impute the missing data at the final follow-up time point. In this case, last observation for the primary endpoint at six months or 12 months will be

carried forward to the 24 months missing data.

For this approach to be valid, there is two underlying assumptions, because the primary endpoint is the composite one, so we should assume there is no adverse event, device failures or neurological failure between the last follow-up and 24 months post-implantation. And also, we would assume that ODI score changed a little, at least improved from six months to 24 months post-implantation.

To assess such assumptions, here I present a table for those, all the completers in both groups. There is a high percentage of completers, more than 70 percent in both groups, who have maintained the success status from the previous follow-up time at six or 12 months.

Since ODI score is a major reason for the device, for the individual overall failure at 24 months, and it's a major dominant reason for the observed difference between the success rate of the two groups at 24 months, I'm going to take some time to talk about how ODI score changed over the follow-up time between these two groups.

As a visual summary of ODI score distributions over the whole follow-up period from month 0 to 24 months, this slide shows the box prods of ODI score over the 24 months follow-up with the median values connected by the line. The blue solid line is for Charite and the red box with dotted line is for the control, BAK.

The main message for this slide is that, as you can see, at early follow-up time from baseline to six months, both BAK patients and the Charite have a decreased ODI score, relatively faster compared to the later follow-up period. At six months, the ODI score in the Charite group reached, plateaued and maintained the single level through 24 months. In contrast, the BAK patients were continuing to improve in ODI score, i.e. decrease. The small ODI score is better, so the ODI score continued to improve from six months to 12 months for the BAK and reached the plateau at 12 months for the BAK patients.

Therefore, it is reasonable to carry forward the last observation at 12 months to 24 months for both groups, because they all reached the plateau

at 12 months. But if you carry forward six months follow-up date to 24 months, because the BAK patient continued to improve from six to 12 months, such carry forward will be in favor of the Charite and against BAK.

Here is the detailed comparison between sponsor's LOCF and the FDA's modified conservative LOCF. In the sponsor's LOCF after imputation with LOCF, the success rate for all the non-completers is 57 percent, which is near a lower bound of the 95 percent confidence interval from the completers analysis population.

In contrast, at the sponsor's LOCF, the success rate for the BAK is only 28.5 percent, which is far below the lower bound of the 95 percent confidence interval of the completers indicating a bias against the BAK with this approach. The major reason, as mentioned in previous slide, because the ODI score continued to improve from six months to 24 months for the BAK. As you can see, for these six months to 24 months, LOCF, majority of BAK patients, 10 out of 11, was carried forward as failures.

So in a conservative way, we maintain the 12 to 24 month LOCF same as the sponsors did, but we modified the LOCF from six to 24 months in a very, very conservative way in favor of the BAK, treat majority of them as success, 10, except for one patient who showed neurological deterioration at six months, so we treat this patient as failures.

And also, very conservatively, we treat all the six months to 24 months LOCF for Charite group as failures. With such conservative LOCF the success rate for the non-completers in the Charite is only 39 percent below the lower bound of the 95 percent confidence interval of the success rate among the completers, and we have 71 percent success rate for the non-completers in the BAK, which is more than the upper bound of the 95 percent confidence interval of the success rate among the completers.

So as you can see, such treatment is biased in favor of the BAK against the Charite in a way that such conservative LOCF, the 95 percent confidence interval, ranged from -10 to 9.5 percent since the upper bound still below 10 percent delta

margin, we still can claim the non-inferiority of the Charite compared to BAK.

So far all the sensitivity analysis I presented had not taken into consideration any potential confounding fact of some covariates. As the sponsor and our lead reviewer has pointed out, a couple of important covariates need to be considered such as age, gender, body mass index, base level ODI score, pre-operative activity, the disc level, L4 to S1 or pain medication and investigational site.

A repeat measure analysis was updated to evaluate the covariate adjusted comparison between the two groups. Please, note that in this model we treat all missing data as success, because the BAK group has a higher percentage of missing data at 24 months, so such treatment will be in favor of the BAK. With such conservative repeat measure model adjusting for all the covariates, the odds ratio -- before I get into the details of this odds ratio, I would like to spend some time explaining what odds ratio means in case some of you don't know this.

The odds of success is the property of

success divided by the property of failure, and the odds ratio of Charite over BAK is the odds of success Charite divided by the odds of success BAK. Corresponding to the 10 percent delta margin, the equivalent odds ratio for Charite over BAK is 0.67. So if the upper bound of the one-sided 95 percent confidence interval for the odds ratio is beyond -- I'm sorry, if the lower bound, if the one-sided 95 percent confidence interval for the odds ratio is beyond 0.67, then we can conclude the Charite device is at least as good as the BAK.

As you can see from this slide, over the several follow-up times from six months to 24 months, the lower bound of the 95 percent confidence interval of the odds ratio is beyond 0.67. And overall, the average across the follow-up time is still beyond 0.67. So we can conclude, based on the covariate adjustment analysis, the Charite is non-inferior to the BAK.

All the sponsor's claims of the Charite's superiority compared to the BAK goes back to the second endpoint, such as ODI score, pain Visual Analog

Score, quality of life, disc height or may be the primary endpoint at the earlier time point were based on their own adjusted P-values without any pre-specified plan to control the study-wide type and error rate.

I would also like to point out to demonstrate that Charite device provides a benefit at the earlier time point after implantation than BAK. Time to sustain benefit should be compared between the two groups. And actually, one of the sponsor's analysis for time to sustain the success for the primary endpoint did not show any superiority of Charite over BAK.

So to summarize, the statistical analysis provides evidence that Charite is at least as good as BAK, except for the worst case scenario where you treat all missing data as failures for Charite, but success for the BAK. Please, also note that the sponsor's sensitivity analysis using completers plus discontinued of all the randomized patients were treating missing data as failures in favor of the Charite, thus it may be biased against the control,

BAK group.

So based on the conservative FDA single imputation LOCF, there is actually almost equivalent success rate between the two groups, 61 percent, and the true success rate for the Charite patients can range from 54 percent up to 68 percent, and the true success rate for the BAK patients could range from 50 percent to 70 percent.

With regard to the second endpoints, no formal claim should be made without any multi-facility adjustment to control the study-wide type and error. Please, also note that the adverse event might be under reported in the current earlier submission. Most recent available data including those discontinued, overdue and not yet due patients need to be analyzed and submitted. Thank you very much. Now, I would like to turn the podium over to Dr. Graham.

CHAIRPERSON YASZEMSKI: Thanks very much, Dr. Chu. We'll hear from Dr. Graham now and at the conclusion of Dr. Graham's presentation, we're going to break for lunch.

DR. GRAHAM: Good morning. I'm Dr. Jove

Graham. I'm an engineer and reviewer with the FDA and I have asked to conclude the FDA's presentations this morning by commenting on the testing and evaluation of wear debris for this PMA submission.

Wear debris is an issue that concerns us because materials, even when biocompatible in bulk form, can elicit a different biological response when they are in the form of small particulate debris. Specifically, particles that are smaller than 5 microns in size can be engulfed by a macrophage cell causing macrophage activation and inflammatory response and in other orthopedic devices, this can lead to osteolysis and bone resorption.

This wear debris induced osteolysis is a contributing factor in aseptic loosening of other total joint replacements and is thought to be one of the limiting factors on the lifetimes of those devices. So here with the Charite Artificial Disc, we have two articulating surfaces that are going to be sliding against each other under a compressive load over the entire lifetime of the device. The surfaces are ultra high molecular weight polyethylene against

cobalt chrome, one on the top and one on the bottom.

So under these conditions, we expect that some wear debris will be generated. Our question is does this wear debris pose a risk to the safety and effectiveness of the device?

The sponsor has performed three kinds of testing to address this issue as previously presented and the tests are listed here. The wear testing of their actual device is what establishes how much debris we think will be generated and the wear rate. Then by looking at the particles that are generated during that testing, this tells us what the size and expected shape of the particles are going to be. And then finally, the sponsor has conducted a small animal study using a rabbit to evaluate the biological response to that debris.

I think the sponsor has identified exactly the three questions that need to be asked with respect to this issue and they have identified and carried out the appropriate tests to answer those questions. I think we need to keep in perspective what the results can and cannot tell us. What these results do a very

good job of is thoroughly characterizing the expected wear behavior of this device. The thing to remember though is the other thing that we would like to do with these results would be to take them and compare them to results from another spinal disc replacement that would be in the literature, that would be well-characterized with the long well-understood clinical history, and because this is the first PMA for a spinal disc replacement, we cannot do that at this time. That literature and data is not available.

So the closest thing that we can try and compare the results to would be wear data from other orthopedic joint replacements like total hips and total knees, and that literature is certainly abundant and the sponsor has drawn that comparison. I think we just need to be careful about the statements or the conclusions that we can draw, because a spinal disc joint is very different than a hip joint or a knee joint. The anatomy is different. The geometry, the conformity. We would test them differently. There are different loads and different ranges of motion.

And because of that, there are always

limitations to what kind of conclusions we draw about clinical performance based just on preclinical results. But here, I think we want to be specifically careful about trying to make clinical conclusions by comparing preclinical disc testing results to testing of hips and knee replacements.

Okay. The first testing was wear testing of the sponsor's actual device in a simulator machine. The testing parameters the sponsor used are in very good agreement with the ASTM standard that is currently being developed. I emphasize currently being developed, because not ASTM, not ISO, no one has a wear test simulator method for spinal disc replacement that has been validated yet in the way that we consider hip simulators or knee stimulators to be validated. In order to do that validation, you really have to be able to take the devices that have been in your machine and take devices that have been in the body, look at those surfaces and see if they have the same wear patterns, the same wear behavior.

And at this time, we just don't have those specimens from in the body to make that comparison.

We only have the devices that have been in the machine to look at. So everyone's wear test method, at this point, ASTM's, ISO's and the sponsor's is sort of a best guess at how we think we should best simulate the in-vivo psychologic loading conditions. I think it is a good sign that the sponsor's choices match very well with ASTM's best guess.

That said, there are two small differences between what ASTM suggests and what the sponsor has done. ASTM suggests the static compressive load of 1,200 Newtons. Although, ISO actually suggests a cyclic load and the sponsor has chosen to use a cyclic load, which is probably going to be more physiologically relevant than a static load. And you see the numbers are different, but they are all in the same ballpark.

There is also a difference in the modes of motion that were tested. ASTM suggests testing each of the three axises that is flex extension, lateral bending and axial rotation in sequence or all three simultaneously. The sponsor has chosen to use two of these modes at once and either couple flex extension

with axial rotation or couple lateral bending with axial rotation. I think it is important to do some kind of couple testing, so it is good that that the axises haven't been tested individually.

And for this device, from the polyethylene's point of view, the polyethylene core is round. It is radially symmetrical, so I think from its point of view there is not much difference between flex extension and lateral bending. And the sponsor has also chosen to use the same range of motion in those two directions. So I think this is an appropriate mode of testing for this device.

The results showed an average wear rate of .11 milligrams per million cycles with a small height loss. And the sponsor states that this average wear rate is lower than most reported wear rates for polyethylene hip and knee replacements. That statement is true. I would just add that we don't yet know what wear rates are going to be acceptable and tolerable in the spine until we have more spinal wear data.

There were results of looking at the

particles that were generated during that testing. Most of the particles were described as smooth flakes, very few elongated particles, an average diameter between .2 and 1.5 microns. And one of the key points here, I think, is that the majority of the particles generated were less than 1 micron in size and submicron. The sponsor again says that the particle size range is typical of simulator testing retrievals from other polyethylene joints, and that is true. I would just add that particles of the same size could elicit a different reaction in different parts of the body.

For an example, I think, particle transport that is where do the particles go once they are generated could be different in the different locations because of differences in the anatomy. We don't have a synovial capsule around the disc space. The epidural space is continuous up and down along the length of the spine, and the difference is in things like lymphatic drainage. All of these can contribute to differences in the reaction to the same size particle in different places.

Finally, the small animal rabbit study was conducted to evaluate the biologic response to these particles. Two note to make on the sponsor's methods. A 3 milligram dose of the sterile drug polyethylene particles were implanted. These particles were manufactured by freezing and grinding polyethylene resin, but this is a standard way or a typical way of doing things. It is very hard to collect enough particles from the actual simulator testing to even be able to look at the size and shape let alone try and collect enough to actually implant into the rabbit.

So that change should just be noted, but I think this is a reasonable way of generating the particles for this test. The dose used was 3 milligrams, and if that wear rate of .11 milligrams per million cycle is right, then this dose should represent almost 30 years worth of accumulated debris. And I think that's an appropriately conservative dose. One other difference between the particles that were implanted into the small animal and the particles that were seen in the wear testing is the size range. Particles implanted into the animal were between 1 and

10 microns.

95 percent of those were below 5 microns, and this is important, because, as I said in the beginning, 5 microns is about the threshold size that a particle needs to be below in order to be engulfed by the macrophages. So these particles here were small enough to be engulfed by the macrophages and probably activate the same kind of pathways that smaller particles would have. However, remember that the majority of the particles that their actual device generated were submicron in size. And from what I can tell, none of the particles implanted here were submicron. So there could have been a different degree or a different response had the particles been smaller, but we don't know.

Finally, the results, some of the results of that animal study. The first two here emphasize that the cerebrospinal fluids seemed normal and there were no lesions or neuropathology of the cord. This is important because it emphasizes that the sponsor did not observe any reactions that would be specific to the spinal cord or the nervous system. In the

group that received particles versus the Sham control group, there was a greater amount of epidural fibrosis and an increased level of the cytokine aisle six at the three month time point.

Aisle six is one of the cytokines that has been associated with the osteolysis pathway. However, that level seemed to decrease back down to normal at the six month time point and the sponsor looked for and did not see increases in any of the other cytokines that we associate with the osteolysis pathway. There was a marked infiltration of macrophages with phagocytosis particles described as a chronic macrophage reaction in the epidural fibrous tissue.

The particles could clump together in a glomerate of 50 to 300 microns in dense macrophage clusters were described adjacent to these. However, there was giant cell reaction, no evidence of cellular apoptosis and the sponsor looked for and did not see any particles in the lymph nodes or in the distant organs.

So I will just conclude by summarizing

what we know and what we don't know. The wear testing of this device has demonstrated that the device will generate some wear debris. The wear rate was measured at .11 milligrams per million cycle. The wear debris was mostly submicron with an average diameter between .2 and 1.5 microns. And the small animal study demonstrated that particles of polyethylene implanted into the spinal region could cause epidural fibrosis, a macrophage reaction, a transient percolation of aisle six that went away later, but no reactions were seen specific to the spinal cord, registry of the spinal fluid.

Finally, we should just consider the preclinical testing has done a good job of characterizing the expected wear behavior of this device, but we can't necessarily establish safety and effectiveness of any spinal device just by comparing preclinical results to those from the hip or a knee device. Also, the wear test simulator needs to be compared to implanted retrievals any wear test simulator does in order to validate that that simulator is applying the proper loads to the motions.

And finally, just keep in mind that wear induced osteolysis for other orthopedic devices is a long-term complication. It is probably not going to show up in the first two years of one to two years follow-up and may not become a problem or be observed until 10 or 15 years of follow-up. Thank you very much.

CHAIRPERSON YASZEMSKI: Thanks very much, Dr. Graham. I would like to ask that we hold our questions for FDA until after lunch and that we take a break for lunch now. It is now about 20 minutes or so after 12:00. Let's reconvene at 1:20. Thanks. Let's break for lunch.

UNIDENTIFIED SPEAKER: We're eating on the 8th floor, but this will be secure.

(Whereupon, the meeting was recessed at 12:23 p.m. to reconvene at 1:25 p.m. this same day.)

1

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

1:25 p.m.

CHAIRPERSON YASZEMSKI: We're going to start the Panel discussion as soon as we start and then you are going to go ahead. We're just waiting for Dr. Diaz. Okay. He can show up. Good afternoon. It is now 1:25. I would like to call the meeting back to order. I would like to remind the public again that while this meeting is open for public observation, public attendees may not participate except at the specific request of the Panel.

We will now begin the Panel discussion. Two voting members of this Panel will open this part of the meeting with their remarks. Dr. John Kirkpatrick will give his remarks on the clinical information and Dr. Brent Blumenstein will address statistical evaluation of the study. Then the Panel will have a general discussion after which the Panel will focus their deliberations on the FDA questions. Then there will be a second open public hearing and FDA summation and sponsor summation.

After that, the Panel will conclude their

deliberations and vote on their recommendation concerning this pre-market application. The Panel can ask the sponsor or the FDA questions at any time, so, please, interrupt and ask any time any Panel Member has a question. The first lead Panel reviewer is Dr. Kirkpatrick. Dr. Kirkpatrick?

DR. KIRKPATRICK: Thank you, Fellow Panel Members, sponsor, of course, and then the public. I appreciate the opportunity to review this. I am both humbled and honored to be able to provide this review to you. I'm also a little bit stronger after having carried around that box to do the review itself. I would like to -- Mark, the page up does not do anything.

MARK: Page down.

DR. KIRKPATRICK: Okay. Just to go over some basics about my review method, since this is a first product of its kind, I look to what would make common sense, so what are the goals of disc replacement. Then I wanted to look at general principles as stated in the literature. We'll review how the literature has followed those principles.

We'll review how the PMA followed those principles. We'll expand on an important area that, I think, warrants further consideration. Then we will review the goals once again and then summarize some key issues.

As I did not have the amount of time available to me, I would also ask, as the FDA, the sponsor if they could, please, try and keep track of any areas where I may have missed something in your PMA. There was an extensive amount of data and I have already found one correction that I had to make. So if you find other things that I say that are inaccurate, by all means, please, make me aware of it so I can refocus any further discussion after this.

The goals of disc replacement, of course, are to remove the presumed pain generator, which is thought to be the degenerative disc. We then replace that with a device restoring normal motion to the functional spinal unit. The key aspects of the reason that this should be better than a fusion of a lumbar disc is the fact that it prevents adjacent segment degeneration. Long-term pain relief would then be

better than arthrodesis, because the pain at the adjacent segment does not degenerate, because of the continued motion at the affected segment, and so that's the key focus of why you would do a disc replacement rather than a fusion, at this point.

General principles from the literature, we should have normal unconstrained psychologic motion. We should have anterior column support, normal biomechanics, wear resistance, a stable bone implant interface or osteointegration, biocompatability. The device should be set fail safe. By that, we mean that if it does fail, it does not cause further damage, in other words, damaging other structures or other areas of the body. It should be revisable, meaning you can salvage the situation and it should be monitorable.

How does the literature deal with these issues and how well can they cover those general principles? With preclinical testing, normal unconstrained motion has been demonstrated in multi-segmental flexibility testing of a cadaver model. Motion profiles among all the segments as well as testing the individual segment and then testing before

and after replacement of the disc. Preclinical testing on anterior column support has been poorly addressed in the literature.

Normal biomechanics has also been poorly demonstrated in the literature from the standpoint that we do not know, in particular, how the facets are affected. Wear resistance should be studied. Wear testing should include cyclic loading replicating the load in motion for the region intended. Failure or 50 million cycles is what has been cited in the literature studies that I was able to find. Wear assessment and particle analysis every 10 million cycles is an appropriate interval, according to the literature, and the debris analysis, of course, is a key component as well.

Osteointegration of biocompatability or the host device interactions are important. One should look at local tissue cytokines in response to the disc and/or the debris generated. No where debris should be found in the reticuloendothelial tissues. Ingrowth or fixation over a minimum of 30 percent of the bone implant interface or surface should be

demonstrated and the materials, of course, should be biocompatible.

As far as clinical studies, that's where we get into the "failsafe" issue, and that is that failure should not risk other injury to the body. It should be revisable or salvaged by either revision or fusion. It should be monitorable with clinical outcomes, radiographic outcomes, complications incidents and other issues. In general, clinical studies of the literature talk about indications, comparison groups, and in this case, fusion is used. Could non-operative treatment also be a consideration for a comparison group? Complications, success rates, follow-up intervals and length are all key features of clinical studies.

Complications should include loss of function, especially through subluxation, subsidence or dislocation, but also in the literature they talk about loss of motion. Heterotopic ossification is another complication that is reported in the literature as a concern. Excessive wear, migration or breakage, facet degeneration at the index level,

adjacent segment degeneration and, of course, infection of the device itself.

Efficacy measures have included visual analog scales, region and disease specific validated outcome measures, such as the ODI, prevalence of revision or additional procedures to the index level and then radiographic measures, including motion analysis or osteolysis or other radiographic changes. This PMA should be commended for its extensive report. They did a good job at trying to be comprehensive. They made a significant effort on preclinical studies. They coordinated a rather elaborate multicenter trial, which recognized learning curve.

It was randomized after a learning curve at each center. They followed their patients for two years. They had reasonable patient accounting, although we have already heard from some statistical follow-up they have some that are yet to complete the study should be included. And we want to see how they compare to the literature standard. With mobility testing, the mobility testing that was in the PMA, as far as I could tell, was referring to a published

study. They had a two paragraph summary, which for somebody with an interest in biomechanics, it was difficult for me to get enough information to convince me that unconstrained motion is attained.

Anterior support, I think, they did a very good job in the study and I have no concerns. As far as general biomechanics of the replaced spine, the test methods are not well-defined in the literature and, of course, as I mentioned earlier, the difficulty of finding out whether the facets have normal stresses across them after the disc replacement, there is not a good method for it in the literature yet, but I would have hoped that the PMA sponsor would have tried to address that in some sense, and perhaps they have and can provide that data to us later.

As far as wear, the date they presented was up to 10 million cycles. They used coupled motion in an axial rotation and flexion-extension. These two issues, I think, should be considered a little bit further. The 10 million cycle number is low compared to those in the literature. It is also low with respect to what the intended life of the device is to

be. As far as coupled motion, their selection of axial rotation in a flexion-extension mode or axial rotation with a lateral bending mode presents some problems.

They also indicated that in their specimens that they looked at, they found grooves in the specimens in the line of the direction of motion. I would have to question whether if they did flexion-extension coupled with lateral bending, whether that extra motion trying to come out of the groove would actually cause more wear debris or a different type of wear debris. So that would be one suggestion I would have as far as additional data.

They also looked at submicron debris with their animals or excuse me, they found submicron debris with their wear analysis, but their neurotoxicity data looked at from 1 to 10 micron and my concern echos that of the FDA and that is would a different response occur if they used a larger volume of the submicron particles as opposed to the micron and above particles?

With regard to osteointegration, this is

one thing I had to change. I actually missed the sentence that said in there summary of the osteointegration in the PMA. I missed the sentence that said that they are presenting data or reference data that was not in the actual clinical study, so I do need to emphasize that the osteointegration information that I was able to review in the PMA was a reference study, but it was not one that was relevant to the surface coating of the device that is being presented in the clinical arm.

In an ingrowth model, they did have adequate osteointegration. I don't see any data in the PMA that represents any kind of long-term biologic fixation with the device that they circulated. Cytokines and reticuloendothelial tissues were examined well in the reference study as well as in a subsequent study that was published using the same device in the U.S. literature as compared to the European spine literature. And I think that demonstrates the fact that the wear debris doesn't cause a problem. But I can't really give it a true pass on the osteointegration side.

Failsafe, I think, the device is failsafe based upon the two year study. Failures did not result in device related further injury in the study. I think the difficulties with revising it are more approach related. And then revisability, the fusion was used for failures predominately. They did have a retrieval or two, but I think it is potentially salvageable from the standpoint of what they presented. Again, this has to be limited with a two year follow-up.

Is it monitorable? I think they did a good job in clinical outcomes. They used the Oswestry scale for a lumbar spine, which is appropriate, a visual analog scale, work status, SF-36. I do have questions with regard to the neurologic status, their specific measure of how they could do statistics on the neurologic outcome was difficult for me to understand. To do statistics it would seem a lot easier to have a number scale to be able to determine. The changes in neurologic function seem to be more qualitative, rather than defined in quantitive.

Radiographic monitoring, they did range of

motion studies that I thought were good. Other measures for disc replacement were unclear. The radiolucency, I don't think has been defined in the literature or by the sponsor as how to grade that or determine how much is there. In addition, they did not look at adjacent segment radiographic changes. And as far as complications, I think, they adequately reported them within the limits of their study and the goals defined.

Their indications were clearly defined. The comparison group was clearly defined. The success criteria were defined. The results were found mostly to be comparable to fusion. I do have some additional questions on stratification among different indication groups and whether that would improve our understanding. In their indications groups, they did include people with facet changes at that disc level and combined those with people that did not have facet changes at the index level. And my curiosity would be would those two different groups result in a different outcome long-term?

Their follow-up intervals and length were

well-defined, but I question whether it was adequate length. Their complications they talked about loss of function. I think it was reasonably well-reported. It was poorly reported for range of motion and it may be just that I didn't find all the data easily. Heterotopic ossification, I could not find that incidence well-described in the PMA. Wear was not found, which is a good thing in the clinical study.

And then facet degeneration, I didn't see an indicator of whether that was examined. Adjacent segment degeneration, the same question there. And then infection, I didn't see any device specific infections reported. Of course, they did have wound problems, arrythmia around the wound and that sort of thing.

Overall, if you were to look at a grade card like my daughters bring home to me from school, we would see that the literature passes on motion. A failure, in my opinion, on the materials provided, because the reference was not contained in the materials, I think that reference probably does cover enough to satisfy me, but technically I can't approve

that, because I have not seen the entire reference.

Anterior column, I think the literature fails, but our sponsor did a much better job and I would give them a pass on that. Biomechanics, I have to give a failing grade to both the literature and our sponsor. Wear, I think is an almost pass. I think their technique was great, except for the alteration of trying to do the coupled motion in both lateral bending and flexion-extension, and I do think they should extend the length of their wear testing.

As far as osteointegration, I had to give them a fail. Biocompatibility, I believe, they passed. Failsafe, again, is poorly described in the literature, although, it describes what the problem would be and the same thing for the PMA. Fortunately, neither have shown disastrous secondary consequences from the device failing. Revisability, I think they passed for the length of follow-up. And then monitorable, I think, they could use some help on the radiographs as well as I mentioned the neurologic scale.

I'm sorry, I'm hitting the wrong button.

On the length of follow-up, I think, this warrants a further consideration. A key issue on disc replacements is the fact that again the concept of not fusing, but replacing with a disc, is to both remove the pain generator, but also prevent adjacent segment degeneration. With that as the fundamental concept, we need to look at how frequently do you get adjacent segment degeneration after a fusion?

Two reasonable references in the literature on a lumbar spine reasonably well-controlled found that there are -- excuse me, 35 percent at five years will develop adjacent segment degeneration and that study did include multi-level fusions. And in another study that looked at four years with a single level fusion, they found 17 percent at four years. So putting that together, we need to think how soon will we see adjacent segment disorders to be able to prove that the fundamental goal of a disc replacement is actually being attained.

So is two years adequate, is a key question. We might be able to look at some statistics to kind of predict how many patients at what time

period would be appropriate to see that knowing that the literature has given us some data for four and five years of adjacent segment development. And of course, there are also additional suggestions in the literature on the length of time that would be considered appropriate for a disc follow-up, and most of those in the literature do suggest a five to 10 year pivotal time span to be able to determine whether these are effective devices.

Once again, adjacent segment degeneration, I think, the sponsor has failed to demonstrate the absence of this occurring, even at two years, because I could not find, again, the radiographic data to back this up. If they did do this, I would appreciate their showing me how and pointing out to me the proper pages in their PMA.

Summarizing, there are some key issues to consider where the literature reported 50 million cycles, I think, they need to bring up to that level. Representative range of motion, is that truly near physiologic? That also opens up the other questions of how much motion are we going to accept as a

preservation of function and what criteria we would set for loss of motion. And then finally, the adjacent segment degeneration is there less with the disc than with fusion? I don't think it is demonstrated and I also am concerned that two years is not adequate to demonstrate this.

Thank you very much.

CHAIRPERSON YASZEMSKI: Thanks very much, Dr. Kirkpatrick. We're going to next ask Dr. Blumenstein.

DR. KIRKPATRICK: Excuse me. If I may do one other liberty at this point?

CHAIRPERSON YASZEMSKI: Yes, sir.

DR. KIRKPATRICK: I have prepared a list of items that I think would be opportunity for us to consider suggestions to the sponsor. If I may, I would like to just dispense with these to the Panel and to the sponsor?

CHAIRPERSON YASZEMSKI: Please, do. Thank you. While Dr. Kirkpatrick is doing that, we'll ask Dr. Blumenstein to come up and give us his statistical analysis next. And I will ask the Panel Members

immediately after Dr. Blumenstein's remarks we're going to proceed to a general discussion. Any questions the Panel Members have for either of our two lead presenters, Dr. Kirkpatrick or Dr. Blumenstein, you may ask them or any questions you have of either our sponsor or the FDA, you may ask.

When we get through those general questions, we'll then proceed to individually looking at the specific questions the FDA has asked us to consider and we will go around the table on each of those questions. Dr. Blumenstein, we're ready when you are, sir.

DR. BLUMENSTEIN: So I basically agree with the FDA statistician's review. I especially liked all of the finer analyses to make sure everything is meeting all the assumptions. I don't like the sponsor's analysis and I will tell you why in a minute. It's more in the category of nitpicking, but despite the flaws, the product appears to meet the non-inferiority criteria and my goal here is to identify the single best characterization of the non-inferiority outcome.

I personally hate the term intent-to-treat. I think that the correct term is analysis by arm. However, it is a little bit late for me to be making these objections, because the term intent-to-treat is very pervasive. I also hate the term population when referring to a part of the data to be analyzed. The population is that from which we sample, unless you are a Camp Thornian, and most people in here won't know what that means. But the term population, so when you use the term ITT population, that to me doesn't make sense at all.

The sponsor's definition of the ITT population not only does the term not make sense, but it is incorrect, because it deletes randomized patients. The ITT is analysis by arm and it includes all randomized patients. To modify the definition of ITT or analysis by arm by deleting patients is tantamount to saying someone is only partly dead. The FDA statistician also apparently agrees with me on this.

So I'm going to give you a little course in randomized clinical trials 101. In a randomized

clinical trial, the arms that you create is a partition of the patients enrolled based on some random process. As a result of that, these arms represent patient groups, that is the subsets of the patients enrolled, that are stochastically equivalent. And the primary analysis is to compare the arms with respect to whatever effect measure is being used. You are not comparing the interventions. The primary analysis is therefore an analysis by arm, that is comparing the arms enrolled as randomized.

If there is no intervention difference, then the probability of the type one error is that declared in the planning of the trial and so forth, provided all of the other principles are followed, such as repeated analyses and so forth. And so the analysis by arm compares the arms with respect to the outcome measures as influenced by all arm specific actions. Now, ideally, arm specific actions are related to the intervention only. That is in the nice clean trial, everybody gets the intervention intended and they have an outcome measured and you are able to compare the two arms and then the comparison of the

arms really does relate to the interventions.

But the degree to which the differences reflect intervention differences depends on the purity of the implementation of the intervention, that is if some patients don't get the intended intervention or the patients are dropped out from the analysis and so forth, then you may or may not, by comparing the arms, actually be comparing the interventions. The deletions from the arms, that is the groups of patients, erode the stochastic equivalence and between armed differences when there are deletions, represent a combination of the differences in the interventions that might or might not exist and the differences due to deletions.

So that when you have deletions in the pure conical randomized clinical trial sense, you have eroded, you have introduced a factor that is eroding the stochastic equivalence that you implemented through randomization. And deletions based on post-randomization events are particularly honoris, because they are more likely related to an intervention, that is a patient may drop out because of side effects or

decide not to come back because of side effects or you may have intervention implementation issues that affect the arm.

The primary outcome in the trial should be defined for all possible contingencies. In a dichotomous outcome, that is you have success or no success observed, and this can be defined for all contingencies, and this is what should have been done in this trial. If we had a time-to-event outcome, we could have incomplete follow-up and we can handle that through censoring provided certain other assumptions are met. The qualitative measures, such as things like quality life and other kinds of things of that nature, laboratory values are difficult because missing data has to be imputed or you have to use some other technique to fill in where data are missing.

The exceptional outcomes, I call them EOs here for lack of a better term, for a dichotomy are no success, but no opportunity to observe a failure. In other words, a patient drops out before the two year follow-up is -- before you can measure the two year follow-up, in this case, or something along those

lines. If EOs are equally distributed between the arms and independent of the intervention, then we have a minimal problem and it becomes a random thing that perturbs our trial a little bit and we just keep them in and we hope they are working out.

But an existence of an EO, that is an exceptional outcome, can be due to side effects and when that happens then we have the potential for bias. The conservative reproach is to call the EOs not successes and this preserves the ability to do the analysis by arm. That is all patients included. So the primary effective efficacy analysis here is really a non-inferiority analysis. And, in my opinion, this is the analysis by arm that is a true intent-to-treat with a conservative EO coating, that is patients that aren't observed to have the success or failures.

The protocol, as far as I could tell in the massive materials that I was provided, did not specify how to handle EOs. And then there was some fussing about whether the analysis plan existed prior to the time that the database was actually analyzed and so forth. Whatever. The definitive analysis is

analysis by arm, and it best characterizes the magnitude of the benefit and also to the extent that the trial matches the real world. It would also match the real world in the sense that there are patients who drop out before you can measure success.

So the Type 1 error specified in the protocol is .05 one-sided. Now, some would argue with this and say that really the criterion should have been .025, that is .05 divided by 2, and other parts of the FDA are very, very strict about this, that if you are doing something one-sided, then you are always doing it at .025 one-sided. But that's a controversy we won't get into much here. The FDA apparently in early meetings accepted a one-sided .05 criterion for success here.

Now, however, the FDA believes that delta should be 10 percent instead of what was apparently agreed upon earlier as 15 percent. So we have some drift in the definition of success. We also have, you know, the primary analysis not being cleanly defined. A lower significance level for final analysis should also be considered, because there may have been some

data snooping. And if there was an interim analysis, we would be decreasing the final criterion to just under .05 as declared in the protocol.

For example, .048, something along those lines. And therefore, if we can look at a tighter Type 1 error probability of .025, we could have a conservative indication of the robustness of the data. Now, what I'm going to show you now is similar to the sensitivity analyses that were done both by the sponsor and by the FDA. So there's really four analyses here.

The first has delta at 15 as specified in the protocol and alpha as a one-sided .05, and a true intent-to-treat or analysis by arm. We have those rates of success of 55.6 percent versus 45.5 percent. And, of course, this meets the non-inferiority criterion using the black welder test and also the confidence interval that is a little different than the FDA presented it. I can't remember how the sponsor did it, but the confidence interval doesn't include the -15 percent, which would cause it to be inferior.

The next analysis is just going down to the delta at 10 percent, but using the one-sided .05. Again, quite clearly, the sponsor meets the non-inferiority criterion. The next one is delta 15, one-sided .025, just to get an idea of if you were to go for a stricter criterion for making Type 1 error, you still meet the criterion, because you have the P less than .0001 and 95 percent confidence interval still precluding the -15 percent.

Finally, the strictest case of delta 10 and alpha .025 and so these are the conical analyses, that is analysis by arm, the true intent-to-treat with some sensitivity testing varying the delta and the overall alpha, and it is consistent with the sensitivity testing that was done in other situations where all but the worst case scenario was also indicated, success with respect to the non-inferiority criterion.

Now, I can't help but say this. If I were to have the opportunity to design this trial today, I would sure look hard at a failure time primary endpoint, that is something like failure free

survival. And the definition of failure time would be possible in a revision, time to revision or significant side effects or perhaps a decrease in that score that was used or something like that. The arms could be compared using a log rank test.

The advantages of this kind of a primary outcome would be that it captures time and it handles missing data better, and that's just my own opinion. I wanted to get that out. Any questions?

CHAIRPERSON YASZEMSKI: Thanks, Dr. Blumenstein. I'll ask the Panel if they have any questions now for Dr. Kirkpatrick or Dr. Blumenstein. We'll have, of course, an opportunity to do that throughout the general discussion. If there are none, we can begin the general discussion now. And this is an opportunity for Panel Members to bring up any questions they would like to ask of either each other, the FDA or the sponsor.

And perhaps I can start it off while folks are thinking about it. And I would like to start with a sponsor question. Dr. McAfee, may I address a question to you?

DR. MCAFEE: Sure.

CHAIRPERSON YASZEMSKI: Several people have brought up the issue of revisions and have used words like life threatening and maybe impossible to do, and I would just like your opinion. I mean, the study center, I would submit, contain the most experienced surgeons and Mr. Christianson showed us that the training centers are well-setup and that the expectation would be that surgeons who want to do this for the first time get training. It has been also shown, however, that you surgeons in the study center did have a training effect and there was a time to getting good at this.

And would you think that when a surgeon has gone through the training and started to do this, and then is confronted with her or his first revision, what would be your opinion? Would such a surgeon be ready to do that? Should perhaps the most experienced surgeons, like yourself, at the training centers be available for consultation or to, you know, maybe decide whether they should see the patient? I would just like to hear your thoughts on that on revisions.

DR. MCAFEE: All right. And, please, direct my answer, because I have a lot of different ways I could answer that. I have been dedicated to trying to reduce the incidence of these complications. Honestly, I don't see a difference in a dynamic spacer versus any anterior instrumentation device. And I'm going to go right to the more serious problems, and if you could put up slide 166.

I think it's important to focus on the number of cases that really required an anterior revision and personally, I have never had to redo a Charite from the front, but I have published a series of 28 cages. The title of the article is "Revision Strategies for Failed Interbody Fusion Cages," so that's 28 cases. And Ensor Transfeldt from Minneapolis presented 40 cases along the same lines of failed Interbody Fusion cages. And the fact of the matter is you want to do everything possible to avoid having to go from the front again.

It's nice to say well, we have gone the left anterior retroperitoneal approach and then for the revision, we'll go from the right side or, if it's

L5-S1, you would want to do the revision through a transperitoneal approach anteriorly at L5-S1.

The keys are in the randomized part of the series, there are really only two cases that required a repeat anterior procedure, and I'm going to add on here the Kurtz/Peloza case report that we heard, so that would actually be three cases being redone from the front. And actually, one of my points is we heard the case report, but we never heard what the indications were for anterior revision. That Charite device was totally confined within in the disc space and, personally, I do everything humanly possible to try to salvage that for the safe posterior fusion in Side 2, Pedical screws, posterolateral bone graft and that is how you would revise any Anterior Interbody Fusion cage.

So for the first case up there, it was revised at one month. This was a technical problem. Fortunately, it was able to be revised anteriorly. A smaller Charite was placed three days post-operatively. The second case was 20 months. The Charite had to be removed from the front and this was

revised with the Anterior Interbody Fusion. So that's really three cases out of the total 205 randomized series.

And I can tell you that, personally, I try to track down the revisions, because that's what I'm interested in, and so far in the United States of the 700 cases, there have been 13 that have required anterior revisions and nine of those were able to be revised with the Charite.

So one of the key points is you are highly dependent on a well-trained access surgeon. The Van Ooij's series that I presented, in Europe the surgeons tend to do their own anterior procedures. We use three different access surgeons. Their primary interest is vascular mobilization and being able to deal with the great vessels.

So we go from here to slide 167 and 168 and these are the total series of re-operations of the Charite. And to me, you know, we're going to be arguing about adverse events and what constitutes a real neurologic problem, but to me it's really cut and dried. It's very objective. If a patient goes back

to the operating room, that's a failure. So there are 11 patients. There are some on this slide and then the next slide, 168. And you will see that by far the majority of the problems were able to be successfully salvaged with what should be a routine operation for a spine surgeon, and that is a posterior approach.

If the patient has leg pain, then you use that as an opportunity. I have had two cases like this in my 93 patients. The patient wakes up with more leg pain, so immediately we get a CT myelogram. I honestly didn't see anything compressing the nerve root, but I felt obligated to explore the patient, so you do a posterior approach, decompress the nerve roots and then do a fusion in Side 2 with Pedical screws. So that's 11 patients re-operated on in the series of 205.

CHAIRPERSON YASZEMSKI: All right. Thanks very much, Dr. McAfee. May I go around the table and ask now for general discussion questions by any Panel Members for either FDA or the sponsor.

DR. DIAZ: I would just like to make a comment on that last answer. Being a neurosurgeon, I

tend to be a little bit more purist on the view of approaches, and to me a revision is limited strictly to going back to where the operation was. A salvage operation, which is the Pedical screw, I do not believe is a revision. So I think I am glad to see that you presented the 12 cases with a true anterior revision, because those answer the very question that was asked, and also I think they are in agreement with the European experience, which indicates that they are doable. So even though the cases are potentially threatening, I think the approach is possible.

CHAIRPERSON YASZEMSKI: Thanks very much, Dr. Diaz. I would like to come around the table now and let's just come in clockwise order and I will ask, Dr. Mabrey, have you any general comments to make?

DR. MABREY: Yes, for Mr. Cunningham regarding the retrieved material from the animal model. How did you determine the absence of wear debris?

DR. CUNNINGHAM: The retrieved materials from the animals were based on selecting tissue directly overlying the operative level. So there are

two animal studies. There was a rabbit study and a primate study. Which one are you referring to?

DR. MABREY: The primate study.

DR. CUNNINGHAM: Yes, we collected tissue right over the top of the operative level, this was a six month follow-up, and we assayed it for a variety of cytokines, as well as macrophage activity, and we used both plain and polarized light microscopy to assess any evidence of wear particulate.

DR. MABREY: And did you use an Oil Red O Stain?

DR. CUNNINGHAM: Excuse me?

DR. MABREY: Did you employ an Oil Red O Stain for this determination?

DR. CUNNINGHAM: No, we did not.

DR. FINNEGAN: Actually, don't sit down. Mr. Cunningham, don't sit down. A couple questions. Why did you only take your baboons at six months?

DR. CUNNINGHAM: Well, primate studies, first and foremost, are very expensive. So the six month follow-up we decided was optimal based on our experience with Interbody Fusion cages. These are

typically run at three with six month as our longest follow-up, so that's why it was selected.

DR. FINNEGAN: And secondly, what kind of activity level did they have? Were they caged or were they out?

DR. CUNNINGHAM: Yes, they were individually housed in cages and the primate has a rapid post-operative ambulation. They typically are recovered by the second day post-operatively and are back to normal activities of bouncing around their cages, but they were not group housed.

DR. FINNEGAN: And they were not where they could do a large amount of swinging and jumping?

DR. CUNNINGHAM: No, the cages themselves are kind of a double decker style, so they are about 8 feet in height and 4 feet by 4 feet deep, so they do have the capacity to elevate themselves and then land.

DR. FINNEGAN: And then I have one other question for the company, but I don't think this is one you want.

Cross-link polyethylene was brought up, and is that something that is being considered?

DR. SERHERN: I am Hassan Serhern, DePuy Spine. Actually, we are using ultra high molecular weight 10-20, guard 10-20 grade, which is cross-linked only by sterilization of 2.7 megarad.

DR. FINNEGAN: Okay.

CHAIRPERSON YASZEMSKI: Thanks, Dr. Finnegan. Dr. Kim, have you any general comments?

DR. KIM: I have a question for Dr. McAfee. It's more a theoretical question. An interesting point that was brought up is that we're putting these implants into relatively young people, and I think it's a compelling argument that these implants will need to last about 40 years.

What are your thoughts on that? Do you think they will really last 40 years and, if not, what would be your second treatment for this problem?

DR. MCAFEE: Well, I hope they will last 40 years. I tell my patients to really look at LeMaire data, which is up to 11 years, which is pretty good. There are five different main surgeons in Europe that have long-term experience. Honestly, to talk to the patients, 10 years is pretty good outcome.

In other words, if I can avoid doing a fusion for 10 years, most of them would consider it a success, because you look at Allen Hildebrandt's study, you know, 2.9 percent risk of adjacent segment disease, Etebar and Cahill, the same kind of range, 4 percent annual incidence of adjacent segment disease. And you compare that to over 10 years, it's actually a 25 percent, in other words one in four chance, of having to redo the adjacent level.

So I can be honest. I have looked all over and I cannot find a single study on any motion preserving device, whether it's anterior or posterior, and there honestly is not a study to date that I have been able to identify that does show a motion preserving device reducing incidence of adjacent segment disease.

I do think the motion is physiologic and theoretically, it looks pretty good, but having said all that, if I can give a patient 10 years longevity then most of them will accept that.

CHAIRPERSON YASZEMSKI: Thanks, Dr. McAfee. Thanks, Dr. Kim. Dr. Naidu?

DR. NAIDU: You know, I'll reserve my comments to when we actually consider the specific questions.

CHAIRPERSON YASZEMSKI: Thank you. Dr. Kirkpatrick?

DR. KIRKPATRICK: I would like to ask just a couple of follow-ups to Dr. McAfee. What specific indications would you list for an anterior revision other than what I understand you have said, which is inappropriate sizing of the implant or inappropriate placing of the implant, which would then be revised within a reasonable short post-operative period?

DR. MCAFEE: Okay. I'll try to just think off the cuff, because I'm really looking at any Anterior Interbody Fusion case, but I have had to redo those, for example, for a severe infection. You definitely want to redo that from the front, because with a foreign body, you want to remove that. I would use some type of autograft and then go posteriorly after a week's worth of antibiotics.

The second case would be a patient who has either impingement on a neurologic structure or a

vascular structure, and what I would worry about would be any case of migration, and I can get into answering that, but there's actually only five cases in the whole series where there was migration and only one of those required a re-operation, which was from the front.

So it's really any life threatening compression on a vascular structure or neurologic structure or a severe deep wound infection, and there were no deep wound infections in this series that required an anterior removal.

DR. KIRKPATRICK: My second comment is really to my Panel colleagues. Dr. McAfee did quote two cervical studies when he was talking about adjacent segment degeneration. He did not quote any lumbar studies, and I would remain standing by the data that I presented of 15 to 35 percent, which would actually favor seeing more of it in the early phases of a follow-up study.

And then my other question would be to the sponsor. If you have had a chance to review my 13 items, if I have misrepresented anything that is in

your PMA, I would appreciate, once again, being informed of that. Thanks.

CHAIRPERSON YASZEMSKI: Mr. Christianson, would someone from the sponsor like to make a comment, at this time, or reserve that until later.

MR CHRISTIANSON: Reserve until later.

CHAIRPERSON YASZEMSKI: Thank you. Thanks, Dr. Kirkpatrick. Dr. Blumenstein?

DR. BLUMENSTEIN: I don't have anything to add.

CHAIRPERSON YASZEMSKI: Thank you. Dr. Besser?

DR. BESSER: In one of the preclinical studies, it talked about the fact that the center of rotation for the implanted device wasn't exactly the same as for the spine.

Would someone like to comment?

DR. CUNNINGHAM: I'll take that one. Jack, could you cue?

CHAIRPERSON YASZEMSKI: Excuse me, Mr. Cunningham, just so the transcription says Mr. Cunningham speaking.

DR. CUNNINGHAM: Yes. Jack, could you cue 640 for me, please? Sorry, I was only given 10 minutes during the presentation. I really couldn't go into great depth in the biomechanical study undertaken at our laboratory, but in addition to quantifying the multidirectional flexibility properties of the device, as I only reported the range in motion, we also quantified the center of intervertebral rotation compared to the intact spine. Can we move ahead three? This is the whole lecture and I'll just key in on the main parts. Go ahead, Jack, one more, please. Another one. Yes.

What we did was in addition to the -- while we were doing multidirectional flexibility, we obtained five stepwise flexion-extension radiographs under both the intact Charite, BAK reconstructions and BAK combined with Pedical screws and these are shown here as you go from full extension through full flexion. Next slide.

By taking the full extension and full flexion views superimposed on each other and using the method of perpendicular bisectors, you can quantify

the center of intervertebral rotation. Now, that is in contradistinction to the instantaneous axis of rotation. This is a single point from full extension through full flexion of the intact and then the Charite reconstruction. And then we can schematically represent these as shown to the right. Next slide, next slide.

We have seen this. This happens to be the neutral zone data that I was unable to report, which shows the relative similarity between the intact and the SB versus the other two reconstructions. Next slide, next slide, next slide.

And this is if we were to plot these centers of intervertebral rotation. Now, the green ellipse represents a best fit and this is where all the centers of rotation occurred for eight specimens in the intact condition at the proximal adjacent level and the operative level. So the green ellipses across here are identical. In the case of the SB Charite for both the operative and superior adjacent levels, these were almost superimposable, a little bit higher here into the disc space, but very, very close to the

intact condition.

In the BAK reconstruction, of course, this is a device designed to stabilize the spine, and we would not expect it to move at the operative level, but, in fact, it does have a little bit of motion and it forms an ellipse below the intact condition, and above we see that this pattern becomes a little more diffuse both in the BAK and then when we add Pedical screws.

So directly to answer your question, I think this does, the center of intervertebral rotation is reproduced with the SB Charite based on N8 to the intact condition.

CHAIRPERSON YASZEMSKI: Thanks very much, Mr. Cunningham. Dr. Besser, does that answer your question?

DR. BESSER: Yes, that answers my question. Thank you very much. I also had a question about the axial rotation range of motion. It's hard to imagine getting 25 degrees in one subject, which I think was one single individual's data.

DR. CUNNINGHAM: Maybe that would be for

another loading mode. Axial rotation would be 5 degrees or less. In our studies it's usually 3 to 4 for a single functional spinal unit in the lumbar spine.

UNIDENTIFIED SPEAKER: Flexion-extension.

DR. BESSER: I had thought that 25 degrees was in the axial direction, which was --

DR. CUNNINGHAM: No, that would not be axial rotation.

DR. BESSER: I would wonder how. Thank you.

CHAIRPERSON YASZEMSKI: Okay. Thank you. I would like also to hear from our industry and consumer patient representatives. Ms. Maher, industry representative?

MS. MAHER: I actually have nothing to ask right at this minute, but I will later.

CHAIRPERSON YASZEMSKI: Okay. Thank you. Ms. Luckner?

MS. LUCKNER: I have nothing at this moment.

CHAIRPERSON YASZEMSKI: Thank you. Any

other general comments? And if not, we're going to proceed to the specific FDA questions that they have asked us to consider. Okay. Mr. Melkerson, could we perhaps have those questions up one at a time, so everybody can see them?

There is copies of the questions available in the hallway outside the door if anyone would like their own copy, but we'll put each question up as we're deliberating it. And what we will do for each question is I will ask one Panel Member to lead off the discussion and then we'll go around in a clockwise fashion until everybody has had a chance to address it. While Mr. Melkerson is getting that up, we can go ahead and get started.

The first question is, please, comment on the results of the wear debris testing and particulate analysis. And I will ask Dr. Naidu to lead off with this one.

DR. NAIDU: The sponsors tested particles less than 5 microns and the question, the issue here is is testing the submicron particle important? And I think that submicron particles may be more acutely

inflammatory, but as far as the chronic inflammation picture goes, I don't think there would be that much of a significant difference between particles that are less than 5 microns. I think the sponsor has adequately demonstrated that the phagocytizable particles actually induce chronic inflammation changes, and so I'm not too concerned about that as far as the submicron particles go.

But what concerns me most in some of the slides that have been shown today as far as explanted specimens in polyethylene at 9.5 years, at 10 year retrieval where the polyethylene has completely fragmented catastrophic failure, and from what I understand at least, from 1997 on the sponsor has been using cross-linked, not cross-linked, but 2.7 megarad irradiated ultra high molecular weight polyethylene.

The problem is that at two years, you may not see oxidation changes that are significant like the earlier slides shown by an explanted specimen at 1.6 years, but somehow or the other aging has not been accounted for in any of these sponsor studies. When I asked earlier in the day as far as the mechanical

testing on specimens, polyethylene specimens, it was quite clear that these are all vacuum packed specimens. No mechanical testings were done on any of the aged specimens.

By rending a 2.7 megarad radiation dose, no matter what you do, whether it be it in oxidation, oxidated in a nitrogen atmosphere, you will induce aging. The problem is the lack of the aged data on polyethylene, one must remember that these devices are put in young, active individuals and one expects these devices to last a long time.

And therefore, my concern here is not the particulate debris more so than the eventual catastrophic failing of the polyethylene that is actually serving as a cushion material. I don't think that adequate polymer characterization has been done. I don't think that adequate aging studies, mechanical studies in properly aged specimens have been done. So I'm not sure as to the actual ultra high integrity in this case.

What I'm concerned about is in the slides presented, the brittle nature of the polyethylene as

exposed leads me to believe that, somehow, this ultra high has been degraded and has been transformed into high density polyethylene. And therefore, I'm a little concerned about the longevity of the implant and the polyethylene liner in light of the radiation treatment. But nevertheless, as far as inflammatory debris, I am pretty satisfied with that.

CHAIRPERSON YASZEMSKI: Okay. Thanks very much, Mr. Naidu. Dr. Blumenstein, have you any comment on Question 1?

DR. BLUMENSTEIN: No.

CHAIRPERSON YASZEMSKI: Thank you. Dr. Besser, have you a comment on Question 1?

DR. BESSER: No.

CHAIRPERSON YASZEMSKI: Thank you. Ms. Maher?

MS. MAHER: I would actually like to ask DePuy Spine to respond to Dr. Naidu's comments on the aging.

CHAIRPERSON YASZEMSKI: Okay.

MS. MAHER: Bill?

MS. COURIER: I'm Barbara Courier. I'm a

researcher at Dartmouth College. I am a paid consultant to DePuy Spine and my transportation costs were paid to this meeting. I would like to put up slide 303 if I could, please.

You mentioned that the materials that have been tested were irradiated in vacuum and in nitrogen. That is true. However, the packaging was not the type of barrier package that one would expect for a nitrogen irradiated or vacuum irradiated component of today, and what I will show in this slide is that actually the materials that were aged on the shelf for 18 months and for 29 months, the 18 month in the pink squares and the 29 month in the solid blue line, show some oxidation with time on the shelf. And so the materials that were wear tested that had a shelf time did, indeed, have some oxidation. This packaging has been improved and now will be GVF packaging, approved technology in use in the knee.

DR. NAIDU: Can I ask a question?

CHAIRPERSON YASZEMSKI: Dr. Naidu, of course.

MS. COURIER: Yes.

CHAIRPERSON YASZEMSKI: Go ahead.

DR. NAIDU: Well, you show the key tone groups there, but I'm not concerned about the oxidation as much as the isothermal crystallization that is induced at the chain scission.

MS. COURIER: Yes.

DR. NAIDU: Do you have any calorimetric studies as far as documenting that this is really not aged, that you have not destroyed the ultra high molecular weight polyethylene integrity into a high density at 2.7 megarads, because these are catastrophic failures that you show at explanted specimens. These are not like, you know, co-flow, anything like that.

The thing is do you have any crystallinity studies?

MS. COURIER: The specimen that you are referring to, the 9.5 years, number one, we don't know what the pre-implanted shelf life was. That particular specimen was gamma in air, and so there is a potential that it could have up to a six year shelf life prior to implantation and that is a piece of

information that we're trying to obtain to determine what the shelf life was prior to implantation.

But given the fact that it may have had a substantial shelf life before implantation, the fact that it showed fatigue failure in-vivo should come as really no surprise and that crystallinity would be extremely high. It would no longer be characterized as an ultra high molecular weight polyethylene.

DR. NAIDU: Can I ask another question? I'm sorry to take up time, but the thing is whether you gamma radiate in air or not.

DR. GAINES: Excuse me.

DR. NAIDU: Okay.

DR. GAINES: Mark Gaines, DePuy Orthopedics, if I could make a comment. The packaging has been changed to GVF, which eliminates all on-shelf oxidation. That is our material of choice currently for our knee product line and has been since 1969. We have done extensive wear testing on that material and we have done accelerated aging and wear testing, accelerated aging with harsh conditions, five atmospheres of oxygen, 70 degrees centigrade for 14

days, which simulates a very severe oxidation condition. And although we see some elevated wear rates, we do not see delamination problems with that and we do not see fracture problems with that material with wear studies that have gone out on a knee simulator to 8, 9 million cycles.

DR. NAIDU: So you do have crystallinity data on these, on the aged specimens? What I'm talking about is not oxidation phenomenon itself. I'm talking about the chain scission that is induced that leads to crystallization no matter whether in the presence of oxygen or not. I'm talking about the integrity change in the ultra high itself. So you do have some crystallinity data that is not presented. Is that what you're telling me?

DR. GAINES: I do not have any here, but we have measured that, yes.

DR. NAIDU: Okay.

DR. GAINES: Yes.

DR. NAIDU: All right. Thanks.

CHAIRPERSON YASZEMSKI: Thank you very much. Ms. Luckner?

MS. LUCKNER: No.

CHAIRPERSON YASZEMSKI: Dr. Witten?

DR. WITTEN: Nothing to add.

CHAIRPERSON YASZEMSKI: Thank you. Dr. Diaz?

DR. DIAZ: Nothing to add.

CHAIRPERSON YASZEMSKI: Dr. Mabrey?

DR. MABREY: I guess after having seen this device and held it in my hands and also looked at Dr. Kurtz' presentation, too, I have to wonder if we really are dealing with a new type of joint. Whether or not there is an actual synovial capsule around it or not, you have two moving surfaces over poly that gets surrounded by scar tissue or fibrous tissue and I think, you know, if we go back to slide 2 in Dr. Kurtz' presentation you can see that that material gets pumped into all those little crevices.

My concern is that six months or a year or even two years may not be long enough to look at the effects of the smaller particulate debris. I can appreciate that there was no evidence of cytokine activity around the explanted material, but I would be

very interested in seeing results from the explanted revisions.

I know it's not always fair to ask people to characterize the tissues around one's failures, because that certainly doesn't look at the majority of your successes, but, nonetheless, I think looking at the tissue if that's available from those devices that have been explanted would be very helpful in characterizing the particles, and I do think that the smaller particles may be a problem in the longer run.

I think over two years it's not a problem, but at least in the total joint realm, we usually don't see evidence of osteolysis until about 36 months or later. So we're looking at a longer time frame now to look at the effects of osteolysis, and I think we need to be aware of that. It wasn't necessarily a question, and it's not actually addressed to any one individual, but it's just something that we have to keep in mind.

I also wonder if we could estimate the total number of particles in those retrieved specimens. I can appreciate the material from Dr.

McKellip's and Dr. Campbell's lab. I know them very well. And you reported on the results from each specimen, but I think we need to go one step further and calculate the total amount of material that is in the retrieved material. I'm sorry, the total amount of wear debris that is within the retrieved material.

CHAIRPERSON YASZEMSKI: Thanks very much, Dr. Mabrey. Dr. Finnegan?

DR. FINNEGAN: I guess mainly a comment, perhaps a question, and I don't mean to sound as scary as I'm probably going to sound, but this has got to be the first time I have seen spine surgeons talk calmly about epidural fibrosis and chronic inflammation, and my concern is that nerve tissue appears to have some long-term response to chronic inflammation and certainly in the brain, amyloidosis appears to be a problem.

So my question is have you done any cell culture studies with nerve tissue with chronic inflammation and have you, in fact, done any correlation with the amyloid literature to see if, in fact, there are any concerns?

MR. CHRISTIANSON: Bill Christianson from DePuy Spine. I have checked with my colleagues who are not aware of any studies that any of us have performed looking for the factors that you just mentioned.

CHAIRPERSON YASZEMSKI: Thank you, Mr. Christianson. Dr. Kim?

DR. KIM: I have nothing to add.

CHAIRPERSON YASZEMSKI: Thank you, Dr. Kim. Dr. Witten, we have gone around the table and discussed wear debris and particulates. In general, the Panel thought that the testing done by the sponsor has been adequate. There were several concerns. These included a request for perhaps considering data on aged specimens. The sponsor has indicated that the same material that they use for this PMA device is a material that they have used for a long time in their total joint replacements and have, in fact, done some of that data, some of those studies, excuse me.

Dr. Mabrey brought up that perhaps, although the disc is a synthesis, it may turn into a synovial like joint after being excised and having the

device encapsulated, and cautioned us that we may need to look for a longer time to really test whether the particulates are going to have an effect and maybe the wear data needs to be done perhaps for 50 million cycles.

And Dr. Finnegan brought up that the neural tissues do seem to have a peculiar response to inflammation and no particular studies have been done to address that question.

Have we adequately discussed Question 1 from the FDA's perspective?

DR. WITTEN: Yes, thank you.

CHAIRPERSON YASZEMSKI: Thank you, Dr. Witten. We're going to move on now to Question 2. If I might ask to have advanced Question 2, asks if there is a higher incidence of the following adverse events occurred in the Charite group compared to the BAK group. These were non-device related pain, wound infections and device related additional surgery at the index level.

We have been asked to discuss the clinical significance of these and any other adverse events

seen in the trials, so this question is the clinical significance of adverse events. Once again, we'll move in a clockwise direction and this time we'll begin with Dr. Kim. Excuse me, Dr. Kim, I'm sorry. Dr. Mabrey, let's start with Dr. Mabrey this time and move around.

DR. MABREY: Thanks. I guess as far as the clinical significance of the differences in those incidents of pain and infection, the first thing we have to realize is we're not comparing apples with apples. I mean, this is a moving device. It has a slightly different micro environment around it compared to the fusion cages, number one. But I would point out that the non-device related pain complications were, it appeared to be, twice as great with the Charite device compared with the BAK, that the infections appear to be double that of the BAK device, although these did not appear to be device related and that additional surgery related to the device appeared to be at a rate of about four times that of the BAK.

I understand that it is a moving device

and it's more prone to failure and that it may not be fair to say that something is four times the rate when you're looking at 3.9 percent versus .9 percent, but those are the figures that I was presented with.

I guess I would ask one of the clinicians if you could comment on the infections. These were all non-device related, meaning they did not appear to originate at the disc space. Is that correct?

DR. BLUMENTHAL: Slide 130, please. Scott Blumenthal. In discussing this question, a few things that we have to keep in mind. Number one is the way that the study was performed, the incidence of reporting AEs was exquisitely sensitive as it should be. Of the three bullet points in terms of non-device related pain infections and device related additional surgery, as mentioned, the numbers were not great. They did not achieve statistical significance.

In terms of the infections, as mentioned in the presentation, none of these were device infections, so we have no infected total disc replacements or BAKs. If a patient had a minor UTI or some redness around the incision, those were reported

as wound infections whether they documented bacterial growth or not. Why there is a difference between the two groups, again, the numbers were not that great. There is not a clear explanation for that. The next slide, 131, the next slide.

Now, in terms of looking at the non-device related pain, most of these were at early follow-up points, again, not statistically significant. They were transient pain complaints and, again, not device related. And when you look at the overall outcomes, they did not seem to affect the overall outcomes particularly and including patient satisfaction scores.

Finally, the device related additional surgery at the index level, this was an interesting one, because it's really just a matter of reporting and how it was reported. If you add the additional surgeries for pseudoarthroses in the BAK group, some surgeons did not report this as being device related. And if you add those nine cases in, then the numbers equalize a bit more.

DR. MABREY: I would like to compliment

the investigators on being honest enough to report the spider bite, 685 days out of surgery.

CHAIRPERSON YASZEMSKI: Thanks very much, Dr. Mabrey. Dr. Finnegan?

DR. FINNEGAN: No comment.

CHAIRPERSON YASZEMSKI: Thanks, Dr. Finnegan. Dr. Kim?

DR. KIM: I do want to echo also that the complication rate is surprisingly low and I'm impressed at how low they both are.

CHAIRPERSON YASZEMSKI: Thank you. Dr. Naidu?

DR. NAIDU: No further comment.

CHAIRPERSON YASZEMSKI: Thanks. Dr. Kirkpatrick?

DR. KIRKPATRICK: Just to help the FDA in thinking this through, from the standpoint of infections, even if they had one device related infection, I would not suspect that that's enough to say that the device itself is a problem. It would take thousands of cases of the device to be able to get enough numbers to find a statistically significant

difference, and I think that is an onerous request of the sponsor. So, you know, if they had 10, at this time, yes, that's a major deal. But since they have had no device specific infections, I don't think it's a concern.

The second issue is in thinking about additional surgery at the index level, as a spine surgeon we often do multiple different procedures on the spine. A patient with a herniated disc at age 30 may end up with a fusion at age 50. But when you're doing the herniated disc at age 30, you don't go straight to the fusion. That is because you're trying to maintain as much function as possible for that motion segment.

This is another step in the anarchy between a basic spinal problem and actually eliminating the motion. So I think it's appropriate that their number of surgeries at the index level was actually higher than what we would expect for BAK fusion, because we would expect it to fuse and no longer need a procedure at that level unless there is a pseudoarthrosis. So that brings no concern as far

as this question.

CHAIRPERSON YASZEMSKI: Thanks, Dr. Kirkpatrick. Dr. Blumenstein?

DR. BLUMENSTEIN: I have no comments.

CHAIRPERSON YASZEMSKI: Thank you. Dr. Besser?

DR. BESSER: No comments at this time.

CHAIRPERSON YASZEMSKI: Thank you. Ms. Maher?

MS. MAHER: Nothing to add.

CHAIRPERSON YASZEMSKI: Thank you. Ms. Luckner?

MS. LUCKNER: No comment now.

CHAIRPERSON YASZEMSKI: Dr. Diaz?

DR. DIAZ: I just would like to echo the outstanding honesty and wonderful presentation of the review that the sponsor made in regard to the detail analysis that they undertook to assess clinical and clinically relevant data. I think the infections that we see here are really probably more related to the added fussiness that the extra steps that require the implantation of the disc require.

Having done enough ALIFs, there is a certain amount of things you need to do and when you compare that to adding the three extra little pieces to what you're doing, I can see where you would have perhaps a little bit more manipulation. I don't view that as a major concern nor the clinical pain related problems, because these are a difficult group of people, and to get an accurate improvement in pain related complaints is asking too much. So I think from my view of the data, I am happy with what I see.

CHAIRPERSON YASZEMSKI: Thank you, Dr. Diaz. Dr. Witten, with respect to adverse events and the increased frequency of these events in the Charite, in general, the Panel doesn't feel that this is a large issue and, in fact, several Panel Members complimented the sponsor on a very thorough and honest review of those events that were adverse.

So we actually see no problem with this, and ask if we have answered this question to FDA's satisfaction.

DR. WITTEN: Yes, thanks.

CHAIRPERSON YASZEMSKI: Thank you, Dr.

Witten. We'll move on to Question 3 now, Mr. Melkerson. Although the Charite Artificial Disc was highly successful in relieving pain, there were a significant number of patients who did not obtain pain relief. 12 percent had no pain relief or had their pain worsen and an additional 13 percent had only partial pain relief. The etiology of their unrelieved pain is unknown. Please, comment on the interpretation of these findings.

I will start with Dr. Kirkpatrick this time.

DR. KIRKPATRICK: Thank you. In dealing in the field of medicine and in educating residents, for example, we often have to look at their statements of this is the best treatment for a patient or this is the cause of that problem, and ask the resident is that what you think or is that what you know?

Now, in the case of a tibia fracture caused by a bumper of a car, can we say that that was a cause and effect? Yes. In the case of low back pain, we have to say we don't know. It's what we think. So we get back to the rationale of what leads

to a fusion in degenerative disc disease, and that is the thought that the disc is a pain generator. Provocative discography documents that. The disc is then excised and replaced with a fusion or fused from posteriorly. That has been shown in international literature not to make a huge difference, but the patient outcomes are comparable.

It is thought to be slightly better than non-operative treatment for degenerative disc disease and that still is somewhat controversial because of the measures that are being used and that sort of thing. If there is a difference, it doesn't appear great. So in summarizing the basic concepts, we don't know what back pain is caused from. We think it's caused from a painful disc and in fusing it, we're trying to get an improvement of that motion segments not moving and not being a pain generator.

The concept behind a disc replacement is stepwise trying to preserve that motion, because the follow-up to the fusion is why does the patient still hurt if we fused the level? And the follow-up to that, in theory, has been well, it must be the

adjacent segment is wearing out, too. And in many cases among patients that you see clinically, they will get one level fused, three years later they will get a second level fused, because their provocative discography has now moved up another level. So the idea behind the disc replacement is to prevent that sequence of events, and so you don't see people with multiple levels of lumbar fusion trying to chase this disc pain.

So when we're looking at the fundamental concepts, are we able to answer the question, can we stop the pain from getting worse in the future by keeping the motion going? That would be a summary of an overall concept of what's going on. I would suspect that in most circles, people would say that the reason people failed is adjacent segment degeneration or there was a cause of back pain that we don't quite understand.

For example, as I mentioned I believe in my presentation, they did include some people with facet changes at the index level. If the biomechanics is preserved, that means the facets are still getting

loaded. They still may be painful. By the same token, the adjacent segment can do the same thing. So overall, we don't know what the pain generator is. We can't explain why the 25 percent don't have more pain relief than we would expect.

CHAIRPERSON YASZEMSKI: Thank you, Dr. Kirkpatrick. Dr. Blumenstein?

DR. BLUMENSTEIN: I have no comments.

CHAIRPERSON YASZEMSKI: Thank you. Dr. Besser?

DR. BESSER: No comments.

CHAIRPERSON YASZEMSKI: Ms. Maher?

MS. MAHER: No comments.

CHAIRPERSON YASZEMSKI: Ms. Luckner?

MS. LUCKNER: No comments.

CHAIRPERSON YASZEMSKI: Dr. Diaz?

DR. DIAZ: I believe that assessing pain is like trying to pin jello on the wall. It is not exactly an easy thing to do. In dealing with resident education, we often play games with the residents trying to teach them. Like Dr. Kirkpatrick mentioned, one of the questions we often ask is tell me what the

possible reasons for back pain are, and once we listed over 65 reasons for back pain.

So trying to isolate a result based on maintaining or preserving function at a single level joint that has been replaced answers only one of 65 reasons. And so I do not believe that this question really helps us reach the conclusion that we want to get, whether the procedure is safe and effective, because there is no way to answer this question to anybody's satisfaction.

CHAIRPERSON YASZEMSKI: Thank you, Dr. Diaz. Dr. Mabrey?

DR. MABREY: I would just echo Dr. Diaz' comments that it's very difficult to pin down pain in this type of situation, and I think the reason the question comes up is because the investigators have been so extremely thorough about recording everything that happens with their patients that we're going to see this type of data. And I applaud their use of the SF-36 and all the other factors as well.

CHAIRPERSON YASZEMSKI: All right. Thank you, Dr. Mabrey. Dr. Finnegan?

DR. FINNEGAN: I have a question for the sponsor. Did any of the patients who developed significant heterotopic ossification have a change in their pain level and, if so, what was it?

DR. CUNNINGHAM: Bryan Cunningham. Jack, could you pull up 254? Proactively, we evaluated the incidence of heterotopic ossification, correlated both the functional kinematics based on plain film radiographs, as well as VAS and Oswestry, and I have a bar chart here that I can show you, which demonstrates the comparative ranges on heterotopic. 654, please. Not there? 654, I believe.

DR. MCAFEE: I'll try to fill in while we're looking for the slides, but we had an independent evaluator.

CHAIRPERSON YASZEMSKI: May I interrupt and just say this is --

DR. MCAFEE: Sure.

CHAIRPERSON YASZEMSKI: -- Dr. McAfee for the transcriptionist. Go ahead.

DR. MCAFEE: Paul McAfee from the same center. It's a core lab and we proactively wanted to

look at heterotopic ossification and the incidence. So we had an independent evaluator look at the digitized films, Dr. Justin Tortolani, and he presented this as the Spine Arthroplasty Society meeting.

In the overall incidence -- well, first we developed a generic classification for heterotopic. This is actually some of the slides if Bryan could come back up, but the key was based on Brooker and Wills' classification in the hip, we have developed the same kind of thing for the spine. So Class 0 was no heterotopic bone. Class I was bone, extra bone was present, but not in the disc space. Class III, there was extra bone present in the disc space, but it did not interfere with motion and Class IV meant spontaneous arthrodesis.

DR. FINNEGAN: Class III didn't interfere with motion as he is going to show us or didn't --

DR. CUNNINGHAM: Yes, I actually have case examples of each to show you that.

DR. FINNEGAN: Okay.

DR. MCAFEE: But fire ahead.

DR. CUNNINGHAM: If you could go to the next slide. Thank you. Next. So we looked at, as indicated, all the plain film radiographs and quantified the range of motion, as well as how that correlated with VAS and the ODI scores. Next.

We actually looked at over 6,000 x-rays to quantify all this. We had both A/P lateral and flexion and extension films for a total of 276 patients. Next. As indicated, we used the Cobb Method. We quantified range and motion at the operative level. Next.

That was only based on flexion-extension. Importantly, you can't do axial rotation. You would need an RSA method or something like that to determine the rotation. We also quantified segmental translations occurring at the operative level. Next.

As indicated, Paul went through the classes, but just to reiterate, we had a Class 0, that means no ectopic bone present. I, islands of bone that were not within the disc space. A Class II, HO is present, but not affecting range of motion. III, it appears to be affecting range of motion on either

flexion-extension or lateral bending films. And finally a Class of IV, which is ankylosis of the operative level. Next. And these would just be case examples. And this is a coronal section through a baboon functional unit. Next.

We looked at both the ODI, the VAS and the segmental range of motion. Next. At two year follow-up, the overall incidence of HO was 4.3 percent. That's 12 of 276 patients. The distribution, 11 of those at 4-5. We had one at 5-1. In terms of the classification of the 12 cases, four of those were Class I, eight, Class II. We had no classes of III or IV. It was either Class I and II.

In terms of progression, most of the HO was noticed at the six weeks post-operative interval of 42 percent. By three months we had six of 12 and at the six month time interval, one more patient presented. So most of these patients presented by three months post-operatively. Next.

And this just gives you case examples of each HO. This is an HO Class of I showing some small islands of bone lateral to the disc, but, again, on